Bacteriocins of Streptococcus pneumoniae and its response to challenges by antimicrobial peptides

Majchrzykiewicz, Joanna Agnieszka

Chapter 3

Exploring the function and regulation of a putative pneumococcal peptide and its gene cluster in *Streptococcus pneumoniae*

Joanna A. Majchrzykiewicz, Jetta J. E. Bijlsma and Oscar P. Kuipers

In silico **analysis of the genome sequence of *S. pneumoniae* R6 indicated that *SPR0140-0146* genes might form a functional cluster containing a pneumococcal peptide of unknown function with some resemblance to bacteriocins. Hence, we named the peptide PpuA and the gene cluster *ppuRABCDE*. The cluster encodes PpuR, a putative regulator, the PpuA peptide, PpuBC, two CAAX endopeptidases, PpuD, a putative transporter, and PpuE a putative branched-chain amino acid transporter. Here, we show that expression of the *ppuRABCDE* cluster is strictly linked to the concentration of amino acids and peptides in the growth medium, suggesting that the function of the cluster is related to the general nitrogen metabolism of this bacterium. In line with this, we demonstrate that expression of *ppuRABCDE* is under negative control of CodY, a branched-chain amino acid responsive regulator. Moreover, transcriptional studies showed that PpuR is likely a positive regulator of *ppuABCDE*. Transcriptome analysis of a *ppuR* and a *ppuA* mutant revealed that expression of two other, not yet described putative clusters, are influenced by the *ppu* cluster. Thus, they were designated as a peptide responsive cluster, *prcRABCD*, and a transporter of amino acids, *taaBC*. Transcriptional analysis of the *prcRABCD* and *taaBC* promoter regions confirmed that PpuR and the *ppuRABCDE* cluster influenced their expression. Interestingly, expression of both the *prcA* and *taaBC* promoter changed upon addition of nitrogen containing compounds to the medium, which suggests that both clusters, *i.e. prcRABCD* and *taaBC*, might be involved in the nitrogen metabolism of *S. pneumoniae*, as is the *ppuRABCDE* cluster. In conclusion, we revealed that three gene clusters, *i.e. ppu, prc* and *taa*, are most likely involved in nitrogen metabolism in *S. pneumoniae*. In addition, two regulators of these clusters were identified: namely PpuR regulates *ppuRABCDE* expression together with CodY, and influences that of the *prc* and the *taa* cluster, which suggests that they form a novel regulon in this bacterium.**

Introduction

S. pneumoniae is a common inhabitant of the human upper respiratory tract and can cause serious diseases *i.e.* sinusitis, acute otitis media, pneumonia and meningitis. The successful spread from the nasopharynx to a variety of different tissues in the human body requires efficient adaptation to changes in the quality and availability of nutrients, such as amino acids and to environmental stresses generated by the antimicrobial defenses of other bacterial species, *e.g.* bacteriocins, and those of the host. Bacteriocins are small cationic antimicrobial peptides (AMPs) produced by Gram-positive bacteria. In general, they can be divided into four classes according to their biochemical and genetic characteristics. Class I, the lantibiotics, comprises peptides that require several posttranslational modifications to acquire biological activity. Class II, the non-lantibiotics do not require modification for

their antimicrobial activity. Class III consists of large proteins and class IV of cyclic peptides (163,203,529).

Signature-tagged mutagenesis (STM) screens and *in vivo* transcriptome analysis of *S. pneumoniae* identified *SPR0140-0146* as required for invasive diseases in mice (38,191,390). Additionally this cluster was identified, by use of differential fluorescence induction (DFI), during *in vitro* conditions resembling a mouse infection model (319). Because these genes were found to be important for *S. pneumoniae* virulence but their function and regulation are unknown, we decided to further study the *ppuRABCDE* cluster's function and regulation.

The *SPR0140-0146* cluster, named here the pneumococcal peptide of unknown function (*ppuRABCDE*) cluster, seems to consist in *S. pneumoniae* R6 of six genes, namely *ppuR*, *-A*, *-B*, *-C*, *-D*, *-E*. The cluster is likely organized in two transcriptional units, one consisting of the *ppuR* gene encoding a putative transcriptional regulator of the Rgg/GadR/MutR family and a presumed operon containing the other genes of the cluster starting with *ppuA* (Fig. 1A). Since PpuR shows more than 30% amino acid sequence identity to the positive transcriptional regulators of known bacteriocins, namely BhtR of *Streptococcus ratti* (223) and MutR of *Streptococcus mutans* (421-424), and because the amino acid sequence of PpuA (Fig. 1B) possesses several characteristic features of bacteriocin-like peptides such as a GG-processing site, positively charged amino acids and a high (~11.3) pI value of the putatively processed peptide, we hypothesized that this cluster might be involved in production of a bacteriocin-like peptide, namely PpuA.

The *ppuBC* genes encode proteins belonging to the CAAX endopeptidase family, also known as the Abi family (401) that consists of various, prokaryotic and eukaryotic, mostly hypothetical proteins, of which the function is unknown. Members of this family are putative metal-dependent proteases that are likely linked to a protein/peptide modification and/or secretion processes (77,401,478). Examples are the PlnIL proteins of *Lactobacillus plantarum* that are probably involved in maturation, transport and immunity of plantaricin A, a bacteriocin (112) and the SkkI protein, which gives immunity to sakacin 23K (257). The *ppuD* gene encodes a putative transporter of the major facilitator superfamily (MFS) and *ppuE* encodes a putative branched-chain amino acid transport protein with homology to AzlC (25). The MFS transporters are single-polypeptide secondary carriers that occur ubiquitously in prokaryotes and eukaryotes. The MFS proteins transport small molecules in response to chemiosmotic-ion gradients. The MFS consists of at least 34 families involved in transport of, amongst others, sugars, drugs, nitrate, nucleosides, peptides or amino acids (395,456). The *ppuRABCDE* cluster shows reasonable biosynthetic locus similarity to clusters involved in bacteriocin production, namely mutacin II and Bht-B (223,421). The lantibiotic mutacin II is one of the bacteriocins produced by *S. mutans* and Bht-B, a non-lantibiotic, is produced by *S. rattus* (60,73,223,556). Like the *ppuRABCDE* cluster, the biosynthetic gene locus of mutacin II and Bht-B consists of genes divided into two transcriptional units. The *mutR* gene, which encodes a homolog of the transcriptional

regulator of glucosyltransferase G (*rgg*) of *Streptococcus gordonii* (495), is followed by genes that contain the structural bacteriocin gene and genes that encode proteins involved in bacteriocin modification/processing, transport and immunity. In this study, we investigated the regulation of the *ppu* gene cluster in more detail in order to shed some light on its function.

A.

B.

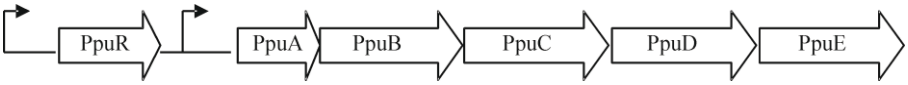

Figure 1. Organization of putative *ppuRABCDE* cluster in *S. pneumoniae* R6 and D39 and amino acid sequence of PpuA. (A) Genetic map of the *ppuRABCDE* cluster in *S. pneumoniae* R6 and D39; thick white arrows indicate the *ppuRABCDE* genes in their transcriptional direction, and two black thin arrows indicate putative promoters of the *ppuRABCDE* cluster, namely P*ppuR* and P*ppuA*. (B) Amino acid sequence of the PpuA peptide; positive amino acids are indicated in bold and the GG-putative processing sites, *i.e.* 1, 2 and 3, are underlined.

Materials and Methods

Bacterial strains and growth conditions

Strains and plasmids used in this study are listed in Table 1. Strains were stored in 10% glycerol (v/v) at -80 °C. *Streptococcus pneumoniae* strains were grown without agitation at 37°C either in liquid media: *i.e.* in M17 (504) (Difco) broth supplemented with 0.5% (w/v) glucose (GM17) and/or Todd-Hewitt (Oxoid) broth supplemented with 0.5% yeast extract (THY) and/or chemically defined medium (CDM) (260), or in solid media: GM17 or THY agar containing 3% defibrinated sheep blood (Johnny Rottier, Kloosterzande, The Netherlands). *Lactococcus lactis* and *Escherichia coli* were grown as described previously (260). When appropriate media were supplemented with antibiotics: chloramphenicol (2 µg/ml for *S. pneumoniae*, 5 µg/ml for *L. lactis*), erythromycin and spectinomycin (for *S. pneumoniae* 0.25 µg/ml and 150 µg/ml, respectively), trimethoprim (18 µg/ml for *S. pneumoniae*), and tetracycline (2.5 µg/ml for *S. pneumoniae*), and ampicillin (100 µg/ml for *E. coli*).

DNA isolation and manipulation

All techniques concerning DNA manipulations were performed as described previously (260,261). Plasmids and primers used in this study are listed in Table 1 and 2, respectively. The chromosomal DNA of *S. pneumoniae* D39 was used as a template for primer design and PCR amplifications. All the constructs were confirmed by sequencing.

Construction of transcriptional *lacZ* fusions

The PP2 plasmid was used to generate a transcriptional fusion of the putative promoter region of *ppuR* to *lacZ*. The putative promoter region was amplified with primer pair PppuR-fv/PppuR-rev. Subsequently, the amplified fragment was digested with *Xba*I and *Eco*RI and cloned into these sites in pPP2 yielding pPP2PppuR. *E. coli* EC1000 was used as a cloning host. Similarly, transcriptional *lacZ*

fusions of the following putative promoter regions of *ppuA* (yielding pPP2PppuA), *prcR* (yielding pPP2PprcR), *prcA* (yielding pPP2PprcA), *taaBC* (yielding pPP2taaBC) and *SPR1352* (yielding pPP2Pspr1352) were constructed with use of primer pairs PppuA-fv/PpuA-rev, PprcR-fv/Pprc-rev, PprcA-fv/PprcA-rev, PtaaBC-fv/PtaaBC-rev and Pspr1352-fv/Pspr1352-rev, respectively. Next the constructs were introduced into *S. pneumoniae* D39 strains by natural transformation.

Table 1. Strains and plasmids used in this study
Ery^R, erythromycin resistance; Tet^R, tetracycline resistance; Spec^R, spectinomycin resistance; trmp^R, trimethoprim resistance

Strain	Description	Reference or source
S. pneumoniae		
D39	Serotype 2 strain, *cps2*	(12,301)
R6	D39(Δ*cps*2 2538-9862) with increased transformation efficiency	(219)
TIGR4		(505,541)
Δ*ppuR*	D39Δ*ppuR*;Spec^R	T. G. Kloosterman
Δ*ppuA*	D39Δ*ppuA*;Ery^R	This work
WH101	D39Δ*codY*;Trmp^R	(199)
Δ*codY*Δ*ppuR*	WH101 Δ*ppuR*	This work
P*ppuR*	D39Δ*bgaA::*P*ppuR-lacZ*;Tet^R	This work
P*ppuR*_1	D39Δ*bgaA::*P*ppuR*_1-*lacZ*:Tet^R	This work
P*ppuR*_2	D39Δ*bgaA::*P*ppuR*_2-*lacZ*:Tet^R	This work
P*ppuA*	D39Δ*bgaA::*P*ppuA-lacZ*;Tet^R	This work
P*ppuA*_1	D39Δ*bgaA::*P*ppuA*_1-*lacZ*;Tet^R	This work
P*ppuA*_2	D39Δ*bgaA::*P*ppuA*_2-*lacZ*;Tet^R	This work
P*prcA*	D39Δ*bgaA::*P*prcA-lacZ*;Tet^R	This work
P*prcR*	D39Δ*bgaA::*P*prcR-lacZ*;Tet^R	This work
P*taaBC*	D39Δ*bgaA::*P*taaBC-lacZ*;Tet^R	This work
P*SPR1352*	D39Δ*bgaA::*P*SPR1352-lacZ*;Tet^R	This work
Δ*ppuR*/P*ppuR*	Δ*ppuR* Δ*bgaA::*P*ppuR-lacZ*;Tet^R	This work
Δ*ppuR*/P*ppuR*_1	Δ*ppuR* Δ*bgaA::*P*ppuR*_1-*lacZ*;Tet^R	This work
Δ*ppuR*/P*ppuR*_2	Δ*ppuR* Δ*bgaA::*P*ppuR*_2-*lacZ*;Tet^R	This work
Δ*ppuR*/P*ppuA*	Δ*ppuR* Δ*bgaA::*P*ppuA-lacZ*;Tet^R	This work
Δ*ppuR*/P*ppuA*_1	Δ*ppuR* Δ*bgaA::*P*ppuA*_1-*lacZ*;Tet^R	This work
Δ*ppuR*/P*prcR*	Δ*ppuR* Δ*bgaA::*P*prcR-lacZ*;Tet^R	This work
Δ*ppuR*/P*prcA*	Δ*ppuR* Δ*bgaA::*P*prcA-lacZ*;Tet^R	This work
Δ*ppuR*/P*taaBC*	Δ*ppuR* Δ*bgaA::*P*taaBC-lacZ*;Tet^R	This work
Δ*ppuR*/P*SPR1352*	Δ*ppuR* Δ*bgaA::*P*SPR1352-lacZ*;Tet^R	This work
Δ*ppuR*/P*ppuA*_2	Δ*ppuR* Δ*bgaA::*P*ppuA*_2-*lacZ*;Tet^R	This work
Δ*ppuA*/P*ppuR*	Δ*ppuA* Δ*bgaA::*P*ppuR-lacZ*;Tet^R	This work
Δ*ppuA*/P*ppuR*_1	Δ*ppuA* Δ*bgaA::*P*ppuR*_1-*lacZ*;Tet^R	This work
Δ*ppuA*/P*ppuR*_2	Δ*ppuA* Δ*bgaA::*P*ppuR*_2-*lacZ*;Tet^R	This work
Δ*ppuA*/P*ppuA*	Δ*ppuA* Δ*bgaA::*P*ppuR-lacZ*;Tet^R	This work
Δ*ppuA*/P*ppuA*_1	Δ*ppuA* Δ*bgaA::*P*ppuA*_1-*lacZ*;Tet^R	This work
Δ*ppuA*/P*ppuA*_1	Δ*ppuA* Δ*bgaA::*P*ppuA*_2-*lacZ*;Tet^R	This work
Δ*ppuA*/P*prcR*	Δ*ppuA* Δ*bgaA::*P*prcR-lacZ*;Tet^R	This work
Δ*ppuA*/P*prcA*	Δ*ppuA* Δ*bgaA::*P*prcA-lacZ*;Tet^R	This work
Δ*ppuA*/P*taaBC*	Δ*ppuA* Δ*bgaA::*P*taaBC-lacZ*;Tet^R	This work
Δ*ppuA*/P*SPR1352*	Δ*ppuA* Δ*bgaA::*P*SPR1352-lacZ*;Tet^R	This work
Δ*codY*/ P*ppuR*	WH101 Δ*bgaA::*P*ppuR-lacZ*;Tet^R	This work
Δ*codY*/ P*ppuR*_1	WH101 Δ*bgaA::*P*ppuR*_1-*lacZ*;Tet^R	This work
Δ*codY*/ P*ppuR*_2	WH101 Δ*bgaA::*P*ppuR*_2-*lacZ*;Tet^R	This work
Δ*codY*/ P*ppuA*	WH101 Δ*bgaA::*P*ppuA-lacZ*;Tet^R	This work
Δ*codY*/ P*ppuA*_1	WH101 Δ*bgaA::*P*ppuA*_1-*lacZ*;Tet^R	This work

Δ*codY*/ P*ppuA*_2	WH101 Δ*bgaA*::P*ppuA*_2-*lacZ*;Tet^R	This work
Δ*codY*Δ*ppuR*/ P*ppuA*	Δ*codY*Δ*ppuR* Δ*bgaA*::P*ppuA*-*lacZ*;Tet^R	This work
L. lactis		
NZ9000	MG1363 *ΔpepN*::*nisRK*	(290)
E. coli		
EC1000	Km^R; MC1000 derivative carrying a single copy of the pWV01 *repA* gene in *glgB*	(303)
Plasmid		
pPP2	Amp^R Tet^R; promoter-less *lacZ*. For replacement of *bgaA* (*SPR0565*) with promoter-*lacZ* fusions. Derivative of pPP1.	(175)
pPP2PppuR	pPP2 P*ppuR*-*lacZ*	This work
pPP2PppuA	pPP2 P*ppuA*-*lacZ*	This work
pPP2PppuR_1	pPP2 P*ppuR*_1-*lacZ*	This work
pPP2PppuR_2	pPP2 P*ppuR*_2-*lacZ*	This work
pPP2PppuA_1	pPP2 P*ppuA*_1-*lacZ*	This work
pPP2PppuA_2	pPP2 P*ppuA*_2-*lacZ*	This work
pPP2PprcR	pPP2 P*prcR*-*lacZ*	This work
pPP2PprcA	pPP2 P*prcA*-*lacZ*	This work
pPP2PtaaBC	pPP2 P*taaBC*-*lacZ*	This work
pPP2Pspr1352	pPP2 P*SPR1352*-*lacZ*	This work

Table 2. Oligonucleotide primers used in this study

Name	**Nucleotide sequence (5' to 3'); restriction enzyme sites underlined**	**Restriction site**
KN-ppuR-fv-1	TGCTCTAGACCTTCTTTTGGATTTGGA	-
KN-ppuR-rev-2	CGGGATCCCATCCTACCACCTCCTAGC	-
KN-ppuR-fv-3	GGGGTACCCATCCCTTTTTGAATTGCG	-
KN-ppuR-rev-4	GAAGATCTAACTGGAAACGACCACAC	-
KN-ppuA-fv-1	CCCACTAGCAGAGGAGGATAGCG	-
KN-ppuA-rev-2	GAGATCTAATCGATGCATGCCCACTTCTGCGACCTAGGAT	-
KN-ppuA-fv-3	AGTTATCGGCATAATCGTGGCTCTTATAGGAGATAATAGG	-
KN-ppuA-rev-4	ACACTGAACTTCTGGTCAGC	-
PppuR-fv	CCGGAATTCCCTTCTTTTGGATTTGGAGGA	*Eco*RI
PppuR-rev	GCTCTAGACATCCTACCACCTCCTAGC	*Xba*I
PppuA-fv	CGGAATTCGCCGAGTTGGAGAGGATGTTACG	*Eco*RI
PppuA-rev	GCTCTAGAGGTTGCCTCCTCTAACATCTTGC	*Xba*I
PprcA-fv	CGGAATTCTGTCCATAATCCCATCTCATAT	*Eco*RI
PprcA-rev	GCTCTAGATAGTTCCAACAGCACTTATCATT	*Xba*I
PprcR-fv	CGGAATTCAGTAGTTCCAACAGCACTTATC	*Eco*RI
PprcR-rev	GCTCTAGATGTCCATAATCCCATCTCATAT	*Xba*I
PtaaBC-fv	CGGAATTCGTAGAAAATGGAACCGTTAAGCA	*Eco*RI
PtaaBC-rev	GCTCTAGATCTGATGCTAAAATCGTTGTAAC	*Xba*I
Pspr1352-fv	CGGAATTCCAACTCCTCCAAGTGATGTGTTGA	*Eco*RI
Pspr1352-rev	GCTCTAGATTGTTCCATGAGATTACCTCGC	*Xba*I
PppuA_S1-fv	ATTTTTTAAAATAAGCCAATTTTCGTGTTATACTG	-
PppuA_S1-2-rev	CGGCGAATTCGCAGGTACCGATGCAT	-
PppuA_S2-fv	CTTCATCTATTATATTCCTCCTTGTTAGT	-
PppuR_S1-fv	CCAAAGTGTCAGAATGTTTTGACA	-
PppuR_S1-2-rev	GGGAAGACAATATCCTCCAAATCC	-
PppuR_S2-fv	GTAGGTTCTTTGTAACCGCTCC	-

In order to construct subclones, shorter pieces of the *ppuR* and *ppuA* promoter regions and the round PCR method with 5' phosphorylated primers was used as described earlier (439). Amplified fragments with primer pair PppuA_S1-fv/PppuA_S1-2-rev yielded construct pPP2PppuA_1, PppuA_S2-fv/PppuA_S1-2-rev yielded construct pPP2PppuA_2, PppuR_S1-fv/PppuR_S1-2-rev

resulted in construct pPP2PppuR_1 and PppuR_S2-fv/PppuR_S1-2-rev resulted in construct pPP2PppuR_2. Subsequently, the constructs were introduced into either *S. pneumoniae* D39, Δ*ppuR* or Δ*ppuA* strains by natural transformation as described before (260,418).

Construction of *ppuA* and *ppuR* mutants

To construct the *ppuA* (Δ*ppuA*) mutant, allelic-replacement mutagenesis was used. Shortly, primers KN-ppuA-fv-1/KN-ppuA-rev-2 and KN-ppuA-fv-3/KN-ppuA-rev-4 were used to generate PCR fragments of approximately 600 bps of the left and right flanking regions of *ppuA*. Next the flanking regions were fused to an erythromycin resistance cassette, generated with primer pair Ery-rev/Ery-for from pORI28, by means of overlap extension PCR (485) and the resulting PCR product was transformed to *S. pneumoniae* D39 yielding Δ*ppuA*.

Construction of the *ppuR* mutant (Δ*ppuR*) was performed as follows. The left and right flanking regions of *ppuR* were PCR amplified with primer pairs KN-ppuR-fv-1/KN-ppuR-rev-2 and KN-ppuR-fv-3/KN-ppuR-rev-4, respectively and cloned as *Xba*I/*Bam*HI and *Kpn*I/*Bgl*II fragments in pORI28spec1 (261) using *E. coli* EC1000 as the cloning host. The resulting construct was used as a template for PCR and amplifies a product with primer pair KN-ppuR-fv-1/ KN-ppuR-rev-4, yielding a linear cassette, which was transformed to *S. pneumoniae* D39. Transformants, having replaced the *ppuR* gene with the *speR* gene, were selected with PCR and verified with Southern blotting.

β-galactosidase assay

β-galactosidase assays were performed as described previously (229,261).

Growth studies

Growth of *S. pneumoniae* D39 was performed in 96-well microtiterplates in CDM. The assay was prepared as follows. A culture of *S. pneumoniae* D39 of approximately OD_{600} 0.2 was stored in aliquots at -80°C. For the growth assay, aliquots were thawed, spun down and resuspended in a fresh medium to OD_{600}~0.1, and were applied into microtiterplates to a total volume of 200 µl/well. The microtiterplate was incubated in a GENios (TECAN Benelux) at 37°C and the OD_{600} was measured every 30 min. All the growth studies were performed in triplicate at least.

DNA microarray analyses and transcriptional profiling

By DNA microarray analysis the transcriptome of Δ*ppuR* and Δ*ppuA* was independently compared to the transcriptome of the D39 wild-type. For DNA microarray analysis each of the strains was grown in 3 biological replicates in CDM and cells were harvested at an OD_{600} of ~0.3. DNA microarrays were produced, prepared and analyzed as described before (261,534,535). Differential gene expression with the Bayesian *p*-value < 0.0001 and with a differential expression greater than 2-fold was considered as significantly differentially expressed.

Synthesis of the PpuA peptide

Two putative versions of predicted mature PpuA peptide, namely PpuA1: (GGGGRSGISGWGVPGIYPGWGNQGMSPNRGAFDWTIDLADGLFGRRRR) and PpuA2: (GGRSGISGWGVPGIYPGWGNQGMSPNRGAFDWTIDLADGLFGRRRR), were synthesized by Pepscan Presto via service of ServiceXC B.V., Pepscan's official distributor in the Benelux. The peptides, delivered as crude, were dissolved in DMSO and desalted with 50 mM Tris-HCl of pH 5.5 on Microcon columns (Millipore), and stored in aliquots at -20°C at concentration of 2 mg/ml.

Putative promoters sequence analysis

Motif identification in the putative promoter sequence of *ppuR* and *ppuA* was carried with the Gibbs Motif Sampler, http://bayesweb.wadsworth.org/gibbs/gibbs.html (508) and Motif Sampler

http://homes.esat.kuleuven.be/~thijs/Work/MotifSampler.html (506,507), and Clone Manager from Scientific & Educational Software (Sci-Ed Software).

Results

The *ppuRABCDE* cluster does not seem to produce a bacteriocin-like peptide

We have indicated that a possible function of the *ppuRABCDE* cluster in *S. pneumoniae* is a bacteriocin-like peptide production. In order to determine whether the PpuA peptide possesses bacteriocin activity, various antimicrobial assays both in solid phase, namely patch and agar diffusion assays, and in a liquid phase, *i.e*. dilution and co-culture assay, were performed under a variety of growth conditions including CDM (data not shown). As putative PpuA producer strains, *S. pneumoniae* R6, D39, TIGR4 and *S. pneumoniae* D39 Δ*codY* were examined. As a negative control, *S. pneumoniae* D39 deficient in *ppuA*, Δ*ppuA*, was analyzed. In the assays various bacterial strains, *e.g*. *Lactococcus lactis*, *Micrococcus flavus*, *Moraxella catarrhalis* and *S. pneumoniae* strains/mutants etc., were examined as indicator strains for their sensitivity to the possible PpuA activity. In addition, a potentially sensitive indicator strain was constructed, Δ*ppu*, by deleting *ppuABCDE* and two downstream genes, *SPR0147-0148*, in *S. pneumoniae* D39 (*SPD0144-0149*). This strain lacks the putative immunity genes and thus should be sensitive to any PpuA bacteriocin-like activity. Antimicrobial activity specifically related to the PpuA peptide was not observed against any tested indicator strain or under any tested growth condition (data not shown). In order to isolate PpuA diverse concentration methods/tools, such as TCA, Amicon Stirred Ultrafiltration Cells, FPLC by use of ion-exchange column, and by generation and induction of a Strep-tagged PpuA, were undertaken (data not shown). Nevertheless, no PpuA-like peptide was identified. Thus, we assumed that the isolation/concentration methods were not suitable or sufficient enough to purify PpuA. Hence, to learn whether PpuA has antimicrobial activity, based on the presence of two putative processing sites in the peptide, two versions of PpuA were chemically synthesized, PpuA_1 and PpuA_2 (Fig. 1B). Subsequently, the PpuA_1 and PpuA_2 peptides were tested for possible bacteriocin-like activity in the spot assay and in the dilution assay (data not shown). Both peptides did not show significant antimicrobial activity (MIC>1 mg/ml) against the various indicator strains mentioned above (data not shown). Thus, we propose that PpuA most likely performs another function in *S. pneumoniae*.

The *ppuRABCDE* cluster is highly induced in chemically defined medium (CDM)

In order to study the *ppuRABCDE* cluster, it was necessary to determine under which conditions the genes were expressed. Therefore, chromosomal transcriptional *lacZ* fusions were constructed to the predicted *ppu* promoter regions, namely P*ppuR* and P*ppuA* (Fig 1A). Subsequently the activity of these promoters was studied in various conditions

(Table 3). Of all the conditions tested, *e.g.* BHI, THY, TSB (data not shown), GM17 and CDM, only the latter induced the expression of both promoters (Table 3). Expression of P*ppuR* increased at most 2-fold, whereas activity of P*ppuA* was induced about 20-fold in CDM (Table 3). This suggested that certain compounds in the complex/undefined media such as peptides might repress both P*ppuR* and P*ppuA*. To confirm this hypothesis, the activity of the promoters was tested in CDM supplemented with casitone, and compared to the activity in CDM itself (Table 3).

Table 3. β-galactosidase activity of transcriptionally fused to *lacZ* promoter of *ppuA* and/or *ppuR*, and their derivatives *ppuA*_1, *ppuA*_2 and *ppuR*_1, *ppuR*_2 in the wild type *S. pneumoniae* D39 and its isogenic mutants of *ppuA*, *ppuR* and *codY* grown in grown in GM17, CDM and in CDM supplemented with 3% casitone. The P*ppuA*_1, P*ppuA*_2, P*ppuR*_1 and P*ppuR*_2, are truncated versions of P*ppuA* and P*ppuR*, and schematic overview of the truncations is shown in Figure 3. Miller Units are the averages of at least three independent experiments and the standard deviations are shown in brackets

Strain	Promoter	β-galactosidase activity (Miller Units)		
		GM17	CDM	CDM without a.a. with casitone
D39	P*ppuA*	46 (11)	987 (75)	220 (16)
	P*ppuA*_1	26 (2)	41 (3)	31 (1)
	P*ppuA*_2	62 (9)	1160 (23)	52 (5)
	P*ppuR*	15 (0.6)	32 (0.5)	17 (1)
	P*ppuR*_1	11 (0.9)	10 (0.4)	11 (0.5)
	P*ppuR*_2	8 (0.8)	9 (0.5)	9 (0.9)
Δ*codY*	P*ppuA*	861 (24)	907 (17)	950 (56)
	P*ppuA*_1	ND	ND	ND
	P*ppuA*_2	753 (44)	1083 (49)	716 (2)
	P*ppuR*	112 (10)	94 (8)	118 (20)
	P*ppuR*_1	163 (5)	146 (11)	ND
	P*ppuR*_2	130 (8)	12.7 (3)	ND
Δ*ppuR*	P*ppuA*	24 (0.5)	20 (4).	18 (3)
	P*ppuA*_1	22 (0.9)	33 (1)	26 (3)
	P*ppuA*_2	72 (12)	44 (7)	36 (2)
	P*ppuR*	20 (2)	33 (1)	17 (2)
	P*ppuR*_1	ND	ND	ND
	P*ppuR*_2	ND	ND	ND
Δ*ppuA*	P*ppuA*	44 (3)	817 (60)	200 (19)
	P*ppuA*_1	ND	ND	ND
	P*ppuA*_2	ND	ND	ND
	P*ppuR*	13 (0.2)	27 (4)	16 (0.5)
	P*ppuR*_1	ND	ND	ND
	P*ppuR*_2	ND	ND	ND

In a medium supplemented with casitone, the P*ppuR* activity decreased 2-fold and the P*ppuA* expression was reduced about 5-fold. This indicated that the components of

casitone, amino acids and peptides, caused repression of *ppuR* and *ppuA* transcription in CDM (Table 3).

CodY is a negative regulator of the *ppuRABCDE* cluster

Expression of P*ppuR* and P*ppuA* decreased upon addition of extra amino acids and peptides (*i.e*. casitone) to CDM (Table 3). CodY is a major bacterial regulator responding to the level of branched-chain amino acids in the environment (108,242,473,484). In addition, transcriptome analysis of a *S. pneumoniae codY* mutant suggested that the *ppuRABCDE* cluster might be a part of its regulon (199). To investigate whether CodY is indeed involved in regulation of the *ppuRABCDE* cluster, expression of P*ppuR* and P*ppuA* was examined in a *codY* deficient strain (Table 3). Transcription of both promoters increased significantly, about 8-fold for P*ppuR* and 4-fold for P*ppuA*, in the Δ*codY* strain independently of the type of medium used and the presence of casitone (Table 3). Therefore, we concluded that CodY is a negative regulator of the *ppuRABCDE* cluster.

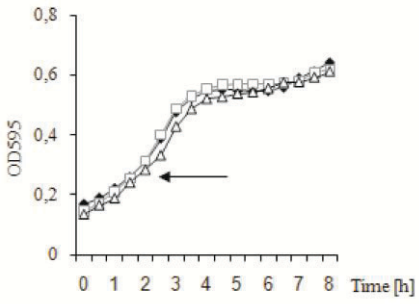

Figure 2. Growth comparison of *S. pneumoniae* D39 and its isogenic mutants in CDM. Comparison of the wild type *S. pneumoniae* D39 (black rhomboids) and Δ*ppuR* (white squares) and Δ*ppuA* (white triangles). The arrow indicates the time point when cells were collected for transcriptome analysis. The bacterial growth curves are representatives of three independent experiments.

Expression profile of the Δ*ppuR* and the Δ*ppuA* strain grown in CDM

Transcriptome analysis of *S. pneumoniae* D39 wild-type and an isogenic *ppuR* mutant was performed, on bacteria grown to an OD_{600}~0.28 in CDM (Fig. 2), to obtain more information about the putative function(s) of *ppuR* and/or the *ppu* cluster in *S. pneumoniae*, and to ascertain whether PpuR is indeed a regulator of the *ppu* cluster. All significant differentially expressed genes can be found in Table 4. The transcriptome analysis showed that in the Δ*ppuR* strain the expression of many genes, particularly hypothetical genes or genes with unknown function had changed significantly (Table 4). The *ppuABCDE* genes were about 10-fold downregulated in the Δ*ppuR* strain, indicating that PpuR most likely acts as a positive regulator of the *ppuRABCDE* cluster.

In addition, we hypothesized that PpuA might act as a pheromone. Therefore, the effect of deletion of *ppuA* on global gene expression was also studied by transcriptome analysis.

Table 4. Summary of transcriptome comparison of *S. pneumoniae* D39 strains: Δ*ppuR* and Δ*ppuA* with D39 wild-type. Only genes, for which the Bayes.P was ≤0.0001 (value between brackets) and transcript levels changed two-fold or more in one or both strains were considered as significantly differentially expressed and are shown below. [a] gene identifiers refer to TIGR4 or R6 locus tags, or to both TIGR4/R6; [b] ratios (D39 Δ*ppuR* compared to D39) in bold, Bayes.P values in parenthesis; [c] ratios (D39 Δ*ppuA* compared to D39) in bold, Bayes.P values in parenthesis; [d] annotation according to NCBI database; NP, not present in TIGR4 genome; NDE, not significantly differentially expressed

Gene ID[a] TIGR4/R6	Gene name	Δ*ppuR*[b]	Δ*ppuA*[c]	Annotation[d]
SP0034		**0.2** (4.2e-12)	**0.2** (5.3e-11)	hypothetical membrane spanning protein
SP0044	*purC*	**0.3** (1.4e-5)	NDE	phosphoribosylaminoimidazole-succinocarboxamide synthase
SP0047	*purM*	**0.4** (2.4e-10)	NDE	phosphoribosylaminoimidazole synthetase
SP0048	*purN*	**0.5** (6.9e-7)	NDE	phosphoribosylglycinamide formyltransferase
SP0051		**0.3** (5.0e-9)	NDE	phosphoribosylamine--glycine ligase
SP0053		**0.4** (4.7e-6)	NDE	phosphoribosylaminoimidazole carboxylase catalytic subunit
SP0054		**0.5** (2.8e-7)	NDE	phosphoribosylaminoimidazole carboxylase ATPase subunit
SP0088		**3.2** (7.9e-4)	NDE	hypothetical protein
SP0090		**0.4** (7.0e-7)	NDE	ABC transporter, permease protein
SP0098		**0.5** (1.7e-6)	**0.5** (1.0e-6)	hypothetical protein
SP0099		**0.5** (7.2e-6)	**0.5** (7.2e-6)	hypothetical protein
SP0100		**0.5** (3.5e-6)	**0.5** (1.9e-4)	hypothetical protein
SP0133		**5.2** (8.0e-4)	NDE	hypothetical protein
SP0138		NDE	**3.2** (6.2e-6)	hypothetical protein
SP0139		NDE	**4.5** (1.3e-8)	UDP-glucose dehydrogenase
SP0141	*ppuR*	**0.1** (2.6e-7)	NDE	transcriptional regulator
SP0142	*ppuA*	**0.0** (9.0e-14)	**0.0** (5.7e-13)	bacteriocin-like peptide
SP0143	*ppuB*	**0.0** (0.0e+0)	**0.0** (0.0e+0)	CAAX amino terminal protease family
SP0144	*ppuC*	**0.0** (1.1e-16)	**0.0** (0.0e+0)	CAAX amino terminal protease family
SP0145	*ppuD*	**0.0** (0.0e+0)	**0.0** (0.0e+0)	transporter, major facilitator family protein
SP0146	*ppuE*	**0.3** (6.9e-06)	**0.2** (9.9e-10)	putative branched-chain amino acid transport protein azlC
SP0159		**0.5** (2.5e-5)	**0.7** (6.4e-10)	homolog of a transporter for Mn(II), Mn(III), and Fe(II)
SP0177		**3.1** (9.7e-5)	NDE	riboflavin synthase subunit alpha
SP0179	*ruvA*	**2.7** (2.1e-5)	NDE	holliday junction DNA helicase motor protein
SP0200		NDE	**2.0** (3.1e-4)	competence-induced protein Ccs4
SP0202	*nrdD*	**2.0** (8.5e-9)	NDE	anaerobic ribonucleoside triphosphate reductase
SP0204		**3.6** (1.1e-6)	**1.9** (6.1e-4)	predicted acetyltransferase, GNAT family
SP0205	*nrdG*	**2.2** (5.8e-5)	NDE	anaerobic ribonucleoside-triphosphate reductase activating protein
SP0206		**7.1** (4.5e-5)	NDE	hypothetical protein; uridine kinase
SP0207		**5.4** (6.3e-4)	NDE	hypothetical protein; uridine kinase
SP0237	*rplQ*	NDE	**2.1** (1.2e-7)	50S ribosomal protein L17
SP0246		**3.1** (6.3e-4)	NDE	DeoR family transcriptional regulator
SP0261		**0.5** (7.7e-9)	**0.4** (7.8e-9)	undecaprenyl diphosphate synthase
SP0281	*pepC*	**0.4** (4.3e-10)	**0.4** (3.2e-7)	aminopeptidase C
SP0303	*celA*	**0.1** (2.5e-12)	**0.1** (2.3e-8)	6-phospho-beta-glucosidase
SP0306		**0.2** (1.1e-10)	**0.1** (2.8e-8)	transcriptional regulator
SP0307		**0.3** (3.5e-4)	**0.2** (1.1e-6)	PTS system, IIA component
SP0309		**0.2** (1.6e-7)	**0.1** (3.9e-7)	hypothetical protein
SP0335	*ftsL*	**0.7** (2.6e-5)	**0.5** (6.5e-7)	Cell division protein FtsL, putative
SP0336	*aliA*	**0.5** (7.1e-8)	**0.5** (9.2e-8)	oligopeptide ABC transporter
SP0338		**2.5** (2.4e-6)	NDE	ATP-dependent Clp protease
SP0355		**0.2** (2.4e-4)	NDE	hypothetical protein
SP0423	*accB*	**0.4** (6e-6)	NDE	acetyl-CoA carboxylase biotin carboxyl carrier protein subunit
SP0437		NDE	0.5 (2.3e-5)	glutamyl-tRNA (Gln) aminotransferase subunit A
SP0449		NDE	2.4 (8.0e-7)	hypothetical protein
SP0506	*vanD*	NDE	2.0 (8.4e-5)	phage integrase family integrase/recombinase
SP0525	*blpS*	**3.1** (9.9e-5)	NDE	regulatory protein
SP0547		**8.1** (8.3e-5)	NDE	CAAX amino terminal protease family
SP0550	*nrrD*	**2.4** (2.1e-4)	NDE	anaerobic ribonucleoside triphosphate reductase
SP0557	*rbfA*	**4.1** (2.9e-5)	NDE	ribosome-binding factor A
SP0595		**4.8** (4.6e-5)	NDE	hypothetical protein
SP0604	*vnc*	NDE	**0.4** (9.7e-7)	sensor histidine kinase VncS
SP0607		NDE	**0.4** (1.3e-7)	amino acids ABC transporter,permease protein
SP0610		**0.2** (1.6e-6)	**0.3** (9.2e-6)	amino acids ABC transporter,ATP
SP0621		**0.3** (1.7e-9)	**0.2** (1.0e-7)	hypothetical protein
SP0634		**3.1** (4.9e-5)	NDE	hypothetical protein

SP0669	*thyA*	**0.4** (9.0e-9)	NDE	thymidylate synthase
SP0670		**0.4** (1.4e-6)	NDE	hypothetical protein
SP0718		**2.2** (5.6e-7)	**2.1** (3.8e-7)	thiamine-phosphate pyrophosphorylase
SP0727	*copY*	**0.3** (3.3e-5)	NDE	CopY rergulator
SP0750	*livH*	**0.4** (2.6e-9)	**0.5** (3.4e-5)	branched-chain amino acid ABC transporter, permease protein
SP0751	*livM*	**0.6** (4.1e-7)	**0.7** (1.9e-4)	branched-chain amino acid ABC transporter, permease protein
SP0766		NDE	**2.0** (2.4e-8)	superoxide dismutase, manganese-dependent
SP0835		NDE	**2.0** (2.2e-5)	purine nucleoside phosphorylase
SP0887		**2.0** (7.4e-4)	NDE	type I restriction-modification system, S subunit, putative
SP0907		**7.5** (4.1e-4)	NDE	hypothetical protein
SP0912		**0.2** (9.2e-10)	**0.3** (2.0e-7)	ABC transporter, permease protein
SP0913		**0.2** (3.7e-7)	**0.3** (2.6e-7)	ABC transporter, ATP protein
SP0921		**0.5** (1.3e-8)	**0.4** (9.2e-10)	agmatine deiminase
SP0943		NDE	**0.5** (7.0e-8)	tRNA (uracil-5)-methyltransferase Gid
SP0968		NDE	**0.5** (2.9e-7)	diacylglycerol kinase
SP1004		**2.4** (6.6e-5)	NDE	conserved hypothetical protein
SP1012		**3.6** (1.3e-11)	**3.7** (3.2e-9)	hypothetical protein
SP1027		**2.0** (5.3e-10)	**1.4** (3.9e-5)	hypothetical protein
SP1041		**2.2** (6.5e-4)	NDE	hypothetical protein
SP1045		NDE	**2.2** (5.6e-4)	hypothetical protein
SP1059		**0.3** (2.8e-4)	NDE	hypothetical protein
SP1137		**0.4** (3.1e-4)	NDE	GTP-binding protein, putative
SP1229		**0.3** (3.5e-12)	NDE	hypothetical protein
SP1320		**0.3** (8.1e-7)	NDE	v-type sodium ATP synthetase, subunit E
SP1325		**0.3** (1.2e-5)	**0.2** (7.4e-5)	Gfo/Idh/MocA family oxidoreductase
SP1342		**2.4** (7.3e-5)	NDE	drug efflux ABC transporter, ATP-binding/permease protein
SP1343		**2.0** (3.6e-5)	NDE	prolyl oligopeptidase family protein
SP1416	*queA*	**0.3** (5.1e-11)	**0.3** (2.5e-12)	S-adenosylmethionine:tRNA ribosyltransferase-isomerase
SP1442		**8.9** (4.6e-4)	NDE	IS66 family Orf2
SP1453		**0.4** (2.2e-6)	NDE	hypothetical protein
SP1460/ SPR1314	*taaB*	**2.1** (1.3e-8)	**2.3** (2.2e-8)	amino acids ABC transporter,ATP
SP1461/ SPR1315	*taaC*	**1.8** (7.4e-9)	**2.0** (9.4e-8)	amino acids ABC transporter,permease
SP1462		**3.0** (3.4e-4)	NDE	hypothetical protein
SP1463	*ogt*	**2.1** (2.7e-6)	**2.1** (8.6e-5)	methylated-DNA--protein-cysteine S-methyltransferase
SP1476		**2.0** (8.0e-4)	NDE	hypothetical protein
SP1499/ SPR1352	*bta*	**1.9** (1.1e-6)	**2.1** (1.0e-6)	bacteriocin transport accessory protein
SP1556		**0.3** (1.1e-5)	NDE	hypothetical protein
SP1588		NDE	**2.0** (3.1e-5)	pyridine nucleotide-disulfide oxidoreductase
SP1589	*murE*	NDE	**0.5** (2.5e-6)	UDP-N-acetylmuramyl tripeptide synthase
SP1704/ SPR1546	*prcD*	**0.1** (4.9e-13)	**0.1** (7.3e-14)	ABC transporter, ATP-binding protein;
SP1705/ SPR1547	*prcC*	**0.0** (1.0e-13)	**0.0** (1.1e-14)	hypothetical protein
SP1706/ SPR1548	*prcB*	**0.0** (1.6e-14)	**0.0** (7.7e-16)	hypothetical protein
NP/*SPR1549*	*prcA*	**0.1** (2.0e-5)	**0.1** (3.2e-5)	hypothetical protein
NP/*SPR1550*	*prcR*	**3.2** (2.7e-3)	NDE	transcriptional activator, Rgg/GadR/MutR family protein
SP1714		NDE	**0.4** (1.5e-8)	transcriptional regulator, GntR family
SP1715		NDE	**0.5** (2.0e-7)	ABC transporter, ATP-binding protein
SP1754		**0.4** (5.4e-7)	**0.4** (8.8e-7)	hypothetical protein
SP1758		**0.2** (7.2e-5)	NDE	glycosyl transferase, group 1
SP1764	*wcaA*	**0.3** (1.3e-5)	NDE	glycosyl transferase family protein
SP1786		**2.1** (1.5e-8)	NDE	hypothetical protein
SP1856		**3.9** (1.1e-4)	NDE	MerR family transcriptional regulator
SP1857		**4.2** (6.2e-4)	NDE	cation efflux system protein
SP1870		**1.6** (4.1e-4)	**2.5** (1.9e-6)	iron-compound;ABC transporter
SP1871		**1.9** (6.0e-5)	**2.9** (4.7e-7)	iron-compound;ABC transporter
SP1872		**2.1** (8.0e-8)	**2.5** (4.1e-8)	iron-compound;ABC transporter
SP1919		**0.3** (6.8e-7)	**0.3** (1.5e-4)	ABC transporter, permease protein
SP1924		**0.5** (1.4e-8)	**0.6** (8.5e-8)	hypothetical protein
SP1925		**0.4** (5.5e-10)	**0.5** (6.6e-9)	hypothetical protein
SP1926		**0.4** (1.2e-10)	**0.5** (1.5e-9)	hypothetical protein
SP2032		**0.2** (8.3e-5)	NDE	BglG family transcriptional regulator
SP2087	*ulaA*	**0.2** (5.6e-7)	NDE	ascorbate-specific PTS system enzyme II
SP2115		**0.4** (2.3e-6)	NDE	hypothetical protein
SP2132		**0.4** (3.7e-9)	NDE	hypothetical protein

SP2160		**0.4** (2.3e-5)	NDE	hypothetical protein
SP2171	*adcC*	**0.5** (1.7e-5)	NDE	Zinc ABC transporter, ATP-binding protein
SP2217	*mreD*	**2.4** (1.4e-6)	NDE	putative rod shape-determining protein

This showed that in the Δ*ppuA* strain the entire putative *ppu* cluster, except for *ppuR*, was about 10-fold downregulated possibly due to a polar effect of the *ppuA* mutation (Table 4). The data suggests that PpuA does not have an influence on *ppuR* expression, which was confirmed by the finding that expression of the P*ppuR* in the Δ*ppuA* strain did not change significantly when compared to the wild type (Table 3).

Comparison of the transcriptome profile of the Δ*ppuR* to that of the Δ*ppuA* strain (Table 4) showed that in both mutants, genes encoding proteins involved in ribonucleotide biosynthesis, cellobiose metabolism, amino acids transporters, iron ABC transporters and many encoding hypothetical proteins had changed expression. Interestingly, in the Δ*ppuA* strain two genes (*SP0138* and *SP0139)* upstream of *ppuR* were induced but not in the Δ*ppuR* mutant, suggesting that PpuA might influence their expression.

Genes *SPR1546-1549* were downregulated nearly 10-fold in both mutants, *i.e.* Δ*ppuA* and Δ*ppuR*. The *SPR1547-1549* genes encode hypothetical proteins. Gene *SPR1546* encodes an ATP-binding protein of an ABC transporter that belongs to a family of ATP-binding proteins of multisubunit transporters involved in drug resistance (BcrA and DrrA), nodulation, lipid transport, and lantibiotic immunity. *In silico* analysis of the genomic region of *SPR1546-1549* showed that these genes might form one transcriptional unit. Interestingly, *spr1550*, which is adjacent to the *SPR1546-1549* genes, was induced 3-fold only in the Δ*ppuR* mutant. The *SPR1550* gene encodes a putative positive regulator of Rgg/GadR/MutR family proteins and *in silico* analysis of the genomic region adjacent to *SPR1546-1550* showed that these genes might form a putative operon. Therefore, the *SPR1546-1550* cluster was selected for further study and we propose the name *prc* for the cluster, which stands for peptide responsive cluster (*prc*) and the proteins encoded *prcRABCD* (*prcR* for *SPR1550* and *prcABCD* for *SPR1546-1549*). Noteworthy, analysis of the genomic region of *prcRABCD* in R6, D39 and TIGR4 showed that TIGR4 lacks *prcR* and *prcA* (*SPR1550* and *SPR1549*, respectively; data not shown).

The *SPR1314-1315* genes, induced approximately 2-fold in both Δ*ppuA* and Δ*ppuR*, encode putative amino acids ABC transporter, thus the name of transporter of amino acids, *taa*, was proposed (*taaBC* for *SPR1314-1315*). To investigate whether *taaBC* responds to Δ*ppuA* and Δ*ppuR* mutation and/or amino acids in the medium we chose it for further study. In addition, *SPR1352* encoding a putative bacteriocin transport accessory protein was approximately 2-fold upregulated in both mutants, namely Δ*ppuA* and Δ*ppuR*, and was chosen for further analysis.

Validation of the microarray results confirmed that PpuR is most likely a positive regulator of *ppuA*

To confirm the observed differential expression patterns of some Δ*ppuR* and/or Δ*ppuA* targets, namely *taaBC, SPR1352, prcR* and *prcABCD*, and *ppuA*, chromosomal transcriptional fusions of *lacZ* with their putative promoters were generated. Expression of P*taaBC* and P*SPR1352* did not change in either the Δ*ppuR* or the Δ*ppuA* mutant (Table 5), in contrast to the transcriptome profiling where 2-fold induction was observed in both mutants. The activity of P*prcR* increased about 4-fold in both mutants, which corresponds to the transcriptome data of Δ*ppuR*, whereas this ORF in the transcriptome analysis of the Δ*ppuA* strain was not significantly differentially expressed (Table 5). Despite the transcriptome results, which showed a 9-fold reduction of *prcA* expression, the transcriptional *lacZ* fusion data showed that P*prcA* decreased roughly 2-fold in Δ*ppuR*.

Table 5. β-galactosidase activity of, transcriptionally fused to *lacZ*, promoter of *taaBC*, *SPR1352*, *prcA* or *prcR* in the wild-type D39 (wt), Δ*ppuR* and Δ*ppuA* strain grown in CDM. Miller Units are the averages of at least three independent experiments and the standard deviations are shown in brackets

Promoter of	β-galactosidase activity (Miller Units)		
	wt	**Δ*ppuR***	**Δ*ppuA***
taaBC	228 (24)	182 (12)	191 (13)
SPR1352	199 (15)	243 (21)	256 (34)
prcA	11700 (2470)	5608 (1440)	192 (17)
prcR	29 (2)	173 (21)	124 (8)

However, in the Δ*ppuA* strain the activity of P*prcA* was reduced nearly 50-fold, which is in agreement with the transcriptome data (Table 5) and might suggest a role of PpuA in gene regulation. In accordance with the transcriptome analysis, P*ppuA* transcription was reduced approximately 45-fold in the *ppuR* mutant (Table 3). However, expression of P*ppuR* was not affected in the Δ*ppuR* strain (Table 3). Thus, PpuR is likely an essential positive regulator of *ppuA* but not of its own expression. What is more, *ppuABCDE* or PpuR likely has a regulatory influence on the *prcRABCD* cluster. These results demonstrate that the activity of P*ppuA* and P*prcA* in general corresponds well with the transcriptome analysis in contrast to the P*taaBC* and P*SPR1352* expression.

Expression of the *ppuRABCDE* putative regulon, *i.e. prc* and *taa*, depends on the presence of nitrogen compounds and/or possibly on the *ppu* gene products

The transcriptome data showed that the mutation of both *ppuR* and *ppuA* influenced the expression of the *prcA* and *prcR* genes (Table 4). However, because in the transcriptional data the effect of the *ppuA* mutation on P*prcA* was stronger than that of *ppuR* (Table 5), we decided to investigate whether this regulation is mediated by the PpuA peptide. Thus, the effect of addition of the synthesized PpuA peptides, namely PpuA_1 and PpuA_2, on the P*prcA* and P*prcR* activity was measured in the wild-type D39, Δ*ppuR* and

Δ*ppuA* strains grown in CDM (Table 6). Because the effect of both peptides was highly similar, only the results of PpuA_1 are shown (Table 6). To determine whether the effects were specific for PpuA_1, this experiment was also performed in CDM with the addition of casitone (Table 6). Expression of P*prcA* and P*prcR* changed significantly in each tested strain upon addition of both PpuA_1 and casitone. In all three strains, expression of P*prcA* was notably higher in CDM than in CDM supplemented with either PpuA_1 or casitone and the effect was most pronounced in the wild type (Table 6). The data suggest firstly that peptides (or di-, tri-peptides, or free amino acids) influence expression of P*prcA* and secondly that either PpuA or product(s) of the *ppu* locus might stimulate *prcA* expression, and thirdly that *prcABCD* might belong to the *ppu* regulon. Expression of P*prcR* in the wild-type strain did not change after addition of PpuA_1 or casitone but increased in both mutants, *i.e.* Δ*ppuR* and Δ*ppuA* with or without PpuA_1 or casitone. This suggests that activity of this promoter is not dependent on a peptide source but rather on the proteins encoded by the *ppu* gene(s).

Table 6. β-galactosidase activity of, transcriptionally fused to *lacZ*, promoter of *prcA*, *prcR*, *taaBC* or *SPR1352* in wild-type D39 (wt), Δ*ppuR* and Δ*ppuA* strain grown in CDM and in CDM supplemented with 10 µg/ml of either PpuA_1 (PpuA_1) or casitone (casitone). Miller Units are the averages of at least three independent experiments and the standard deviations are shown in brackets

Promoter of	Medium	β-galactosidase activity (Miller Units)		
		wt	Δ*ppuR*	Δ*ppuA*
prcA	CDM	11700 (2470)	5608 (1440)	192 (20)
	PpuA_1	4008 (750)	2553 (150)	303 (34)
	casitone	764 (11)	455 (30)	46 (3)
prcR	CDM	29 (2)	173 (11)	124 (8)
	PpuA_1	43 (5)	110 (15)	104 (17)
	casitone	31 (2)	123 (9)	65 (2)
taaBC	CDM	228 (14)	182 (12)	191 (11)
	PpuA_1	530 (24)	170 (18)	108 (5)
	casitone	508 (30)	73 (5)	55 (4)
SPR1352	CDM	199 (17)	243 (33)	256 (24)
	PpuA_1	198 (13)	229 (12)	184 (9)
	casitone	134 (5)	105 (8)	143 (10)

Given the observed effect of peptides (or di-, tri-peptides, or free amino acids) on the P*prcA* promoter, we decided to examine their effect on P*taaBC* and P*SPR1352* expression in the same growth conditions. In the wild-type, the expression of P*taaBC* increased when either PpuA_1 or casitone was added (Table 6). Interestingly, although PpuR and PpuA do not seem to be involved in P*taaBC* expression in CDM, induction of expression in response to either PpuA_1 or casitone did depend on these two proteins (Table 6). The activity of P*SPR1352* did not change significantly in either tested conditions (Table 6). This demonstrates that the *taaBC* genes might be involved in amino acid

transport and that there is a functional and or a regulatory link between the *taaBC* and the *ppuRABCDE* genes.

Prediction of putative PpuR and CodY operators in *S. pneumoniae*

Since CodY is a negative regulator of *ppuR* and possibly of *ppuA*, and PpuR is an essential positive regulator of *ppuA*, this suggests that CodY binding box(es) may be present in the promoter region of at least one of these two genes. In conjunction, there is likely a putative operator site for PpuR, which has not yet been identified, in the promoter of *ppuA*. A previous study on CodY in *S. pneumoniae* showed that except for *ppuR* and *ppuA* the members of the CodY regulon contain a sequence resembling the consensus binding motif of CodY in *L. lactis* (AATTTTGWCAAAATT, CodY binding consensus motif) (108,199). Analysis of the *ppuR* and *ppuA* promoter region, with the Sampler Motif, Gibbs Motif, and Clone Manager and by eye, indicated putative CodY-boxes in both regions (marked as a dashed line in Fig. 3) but no putative operator site for PpuR (Fig. 3).

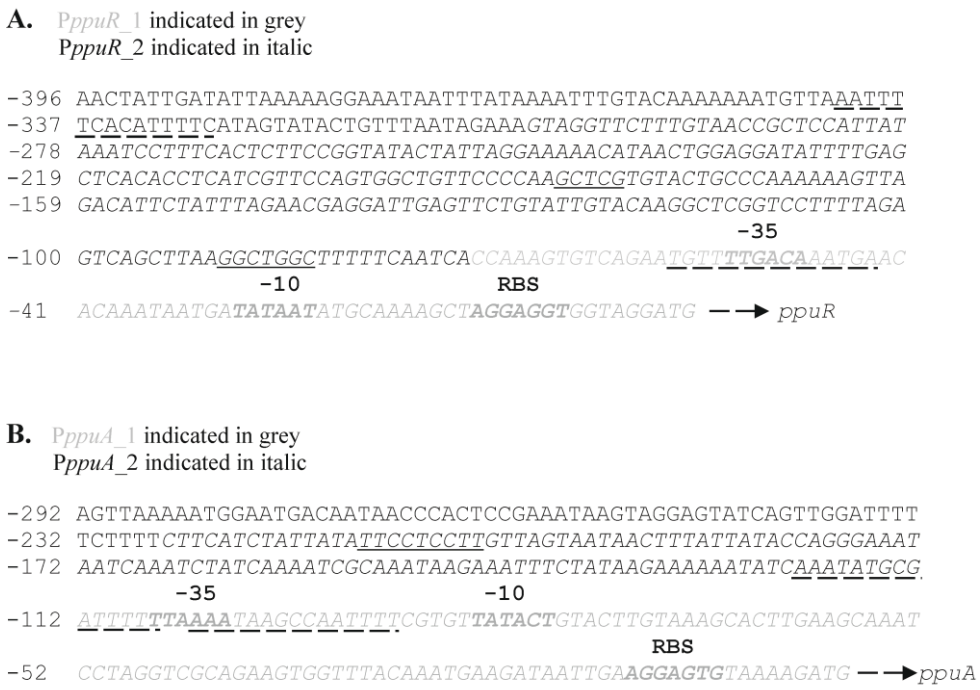

Figure 3. Nucleotide sequences of putative promoter regions of (A) *ppuR* and (B) *ppuA* in *S. pneumoniae* D39. Numbers indicate the base positions relative to the translational start. Predicted -35 and -10 boxes are shown as bolded. Predicted CodY binding motifs are underlined with dashed line. In the *ppuR* and the *ppuA* promoter sequence, underlined bases indicate inverted repeat/palindromic sequence. Designed truncated promoter fragments of *ppuR* and *ppuA*, which were constructed in order to find putative operator sites for PpuR and CodY, are indicated for P*ppuR*_1 and P*ppuA*_1 in grey shade and for P*ppuR*_2 and P*ppuA*_2 indicated in italic.

In order to establish the location of the putative CodY/PpuR operator(s) in the promoter regions of *ppuA* and *ppuR*, each of them was truncated from the 5' end into two shorter fragments, P*ppuA*_1 and P*ppuA*_2 for P*ppuA* and for P*ppuR*, P*ppuR*_1 and P*ppuR*_2, and fused to *lacZ* reporter gene using the pPP2 vector (Fig. 3 A and B). Subsequently, expression of the truncated promoter fragments was measured in various media and genetic backgrounds (Table 3). Activity of P*ppuA*_1 was abolished in all conditions and strains tested, indicating that the PpuR putative operator is not present in this promoter fragment as expression of *ppuA* is strictly dependent on this regulator (Table 3). In contrast P*ppuA*_2 expression was similar to that of the wild type P*ppuA*, in all conditions and strains tested, strongly suggesting that the PpuR putative operator is located in this fragment (Table 3). As all expression of P*ppuA*_1 is abolished it was hard to specify whether a putative CodY operator is present in this promoter. The promoter expression studies of P*ppuR*_1 and P*ppuR*_2 showed that the CodY putative box is most likely located in the P*ppuR*_1 promoter fragment since the CodY repression effect was visible in both truncated promoter fragments and as well in the wild type one (Table 3).

Discussion

The aim of this study was to determine the function(s) and regulation of the *S. pneumoniae* putative pneumococcal peptide of unknown function cluster (*ppuRABCDE*) that has been suggested to be important in invasive disease (38,191,319,390). Interestingly, the *ppuRABCDE* genes were found to be highly induced in blood, indicating their contribution to survival in this environment (390). However, growth in blood of the *ppuA* mutant tested *in vitro* was similar to that of the wild type (390). Genes *ppuR*, *ppuD* and *ppuE* were found in an STM study as important for lung infection (191). Thus, although the functional role of the operon is still unclear it is likely that it plays an important role during pathogenesis.

Since we have shown that CodY, a branched-chain amino acid responsive regulator, is a negative regulator of the *ppuRABCDE* cluster (Table 3), we hypothesize that *ppuRABCDE* is likely to be involved in nitrogen metabolism in *S. pneumoniae*. CodY is a global regulator that adjusts bacterial cell metabolism to the environmental changes in nutrient supply and additionally influences expression of genes involved in virulence (199,417,484). Generally, CodY is activated by branched-chain amino acids (BCAAs) and in *Bacillus subtilis* also by binding GTP (107,125,411,432,469). CodY regulates expression of a broad range of genes, which in *B. subtilis* are involved in transport and metabolism of nutrients, sporulation, motility and competence development (354). In *L. lactis*, CodY represses peptidases, peptides and amino acids uptake systems, and aminotransferases, during growth in complex media (62,164,165). In *Streptococcus pyogenes* and *Staphylococcus aureus* CodY influences, besides genes involved in amino acid transport and metabolism, the expression of genes encoding proteins involved in virulence and

virulence regulation (312,417). In *S. pneumoniae* CodY contributes to colonization of the nasopharynx (199). The putative regulon of CodY in this bacterium includes genes encoding proteins involved in amino acid uptake, metabolism and biosynthesis, and the *ppuRABCDE* cluster (199). Interestingly, the regulon of CodY in *L. lactis* includes genes probably involved in a production of a putative bacteriocin or cell communication peptide (166).

Based on our study we propose the following putative regulation mechanism of the *ppu* cluster (Fig. 4). Expression of the *ppuABCDE* operon is dependent on positive regulation by PpuR, which is repressed by CodY; this regulator might also repress *ppuABCDE*. Interestingly, *ppuR* was constitutively upregulated in a multidrug resistant *S. pneumoniae* M22 strain (321), indicating possible involvement of PpuR in resistance mechanisms in this organism. Notably, the *ppuRABCDE* cluster was up-regulated in D39 Δ*glnAP* grown in GM17 (201). Induction of the cluster is surprising since: *i*) the mutant grew in a rich medium, *i.e.* GM17, *ii*) CodY was likely expressed in this condition and *iii*) this cluster is repressed in the wild type (Table 3). However, it might indicate that glutamine/glutamate deficiency influences regulation of the CodY regulon or of the cluster. Therefore, it would be interesting to verify whether these amino acids affect *ppuRABCDE* expression. Similarly, downregulation of the *ppuABCDE* genes in the *psaR* mutant in D39 strain grown in CDM was unexpected (200), since we showed that the *ppuRABCDE* cluster is expressed in CDM. However, additional regulators such as PsaR might be involved in maintaining *ppuRABCDE* expression.

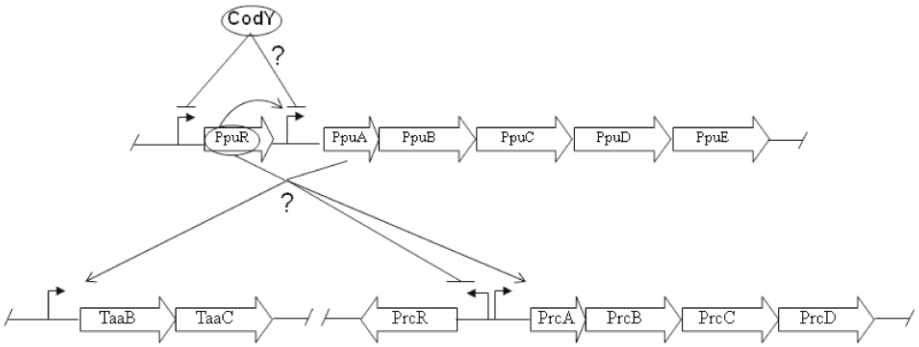

Figure 4. Schematic prediction of regulation of the *ppuRABCDE*, *taaBC* and *prcRABCD* cluster expression. Thick white arrows indicate genes of the clusters. Black thin arrows indicate putative promoters of the clusters. The regulators of the *ppuRABCDE* cluster, *i.e.* PpuR and CodY, are marked in two white circles. PpuR activates *ppuA* expression (open arrow). CodY represses expression of the *ppuR* and possibly (indicated with a question mark) of the *ppuA* gene (perpendicular). A functional or a regulatory link (indicated with a question mark) between the *ppuRABCDE*, *taaBC* and *prcRABCD* clusters are marked with either open arrows or perpendicular.

In silico analysis of the *ppuA* and *ppuR* putative promoter region identified putative CodY motifs in the P*ppuR*_1 fragment (Fig. 3). Whether the box is functional needs to be confirmed and is the subject of ongoing experiments. As with other confirmed CodY boxes, the putative binding sites are present in close vicinity of the -35 box of both promoters (166). However, a previous study in *B. subtilis* demonstrated that even up to five mismatches within the CodY consensus box could result in a functional element indicating that we might have overlooked other authentic CodY boxes (26). It is uncertain whether P*ppuA* harbors a CodY box and whether CodY influences *ppuA* expression directly or only through regulation of PpuR since expression of P*ppuA*_1 was abolished in all mutants probably because this promoter fragment lacks the PpuR operator site (Fig. 3, Table 3). Transcription of P*ppuA*_2 in Δ*ppuR* was approximately 2-fold higher, in all tested media, than that of the intact promoter, which is confusing since the positive regulator, *i.e.* PpuR, was absent in this condition. Thus, more experiments are needed to determine whether CodY directly influences expression of P*ppuA* and this is a subject of ongoing research. A putative operator region(s) of PpuR has not yet been found in P*ppuA*. Hence, to prove a direct regulatory effect of PpuR on the *ppuA* promoter region, direct binding of PpuR protein to this promoter needs to be performed and it is also the subject of ongoing experiments.

Transcriptome analysis of the Δ*ppuA* and Δ*ppuR* strains was performed in order to determine the potential function(s) of PpuA, as well as that of the *ppu* cluster. The response of the Δ*ppuR* and Δ*ppuA* mutant was comparable suggesting that they are in the same pathway (Table 4). In both mutants there were many differentially expressed genes of diverse or unknown function, which made it impossible to pinpoint a putative function for *ppuRABCDE*. The expression of two putative, not yet studied clusters, namely *prcABCD* (*SPR1546-1549*) and *taaBC* (*SPR1314-1315*), decreased and increased, respectively, in both, *ppuA* and *ppuR*, mutants. Notably, gene *prcB* was found in an STM screen as important for lung infection (191) and *prcB* and *prcC* belong to one of the accessory regions (AR) in *S. pneumoniae* (38). The ARs are regions of diversity between *S. pneumoniae* strains and they can have an effect on the ability to colonize and to cause invasive diseases by this bacterium (38). Interestingly, *prcABCD* were among the few genes induced in a *spxR* mutant (431). SpxR is a positive regulator of *spxB* and *strH*, encoding a pyruvate oxidase and a glycoprotein exoglycosidasae, respectively. It is not known, what stimulates the regulatory function of SpxR, but it was hypothesized that perhaps SpxR senses the metabolic state of the cell (431). Notably, SpxR is required for virulence in a murine model of infection diseases (431). Transcriptional studies with the two *prc* promoter regions, *i.e.* P*prcR* and P*prcA*, in Δ*ppuR* and Δ*ppuA* demonstrated an influence of the *ppuRABCDE* cluster on their expression (Table 5). Consequently, transcription of P*prcA* and likely that of P*prcR* might be mediated by Ppu product(s), thus *ppuRABCDE* and *prcRABCD* might belong to the same regulon. Notably, activity of P*prcA* decreased upon supplementation of the growth medium with amino acids and peptides

(Table 6), which might indicate involvement of the *prcRABCD* cluster in controlling nitrogen metabolism in *S. pneumoniae*, as is suggested for the *ppu* cluster. Similarly, because the P*taaBC* activity changed after addition of casitone and/or the PpuA_1 peptide, we propose that the *taaBC* genes encoding putative amino acid transporters are also involved in nitrogen metabolism. What is more, expression of P*taaBC* indicated a functional or a regulatory link between the *taaBC* and *ppu* genes, since the activity of this promoter decreased in both the *ppuA* and the *ppuR* mutant. Expression of the *prc* and the *taa* cluster have not been changed in the transcriptome either of the *codY* mutant (199) or of the Δ*glnAP* mutant, in which expression of the whole CodY regulon was altered (201), indicating that probably they are not directly regulated by CodY and thus they do not belong to the CodY regulon.

All together, we showed that PpuR is most likely a positive, essential regulator of *ppuABCDE* and that CodY is a negative regulator of the *ppuRABCDE* cluster (Fig. 4). Additionally we demonstrated that *ppuRABCDE*, *prcRABCD* and *taaBC* are all possibly involved in controlling nitrogen metabolism in *S. pneumoniae*, which has to be confirmed by further research. Most importantly, these three novel clusters might be linked to each other on a regulatory and/or functionally level and eventually form a regulatory unit.

Acknowledgements

We thank Tomas G. Kloosterman for providing the *S. pneumoniae* D39 Δ*ppuR* mutant and for assistance with the data analyses. We thank Rachel Hamer for her technical help in construction of the P*ppuA* mutant.